SEQUENCE LISTING

<110>   Kwang , Jimmy
         Ling, Ai Ee
         Ooi, Eng Eong
         Chng, Hiok Hee

<120>   Diagnostics for SARS Virus

<130>   2577-162

<150>   60/486,918
<151>   2003-07-15

<150>   PCT/US04/003307
<151>   2004-02-04

<160>   25

<170>   PatentIn version 3.2

<210>   1
<211>   1269
<212>   DNA
<213>   SARS coronavirus

<220>
<221>   CDS
<222>   (1)..(1269)

<400>   1
atg tct gat aat gga ccc caa tca aac caa cgt agt gcc ccc cgc att   48
Met Ser Asp Asn Gly Pro Gln Ser Asn Gln Arg Ser Ala Pro Arg Ile
1   5   10   15

aca ttt ggt gga ccc aca gat tca act gac aat aac cag aat gga gga   96
Thr Phe Gly Gly Pro Thr Asp Ser Thr Asp Asn Asn Gln Asn Gly Gly
            20   25   30

cgc aat ggg gca agg cca aaa cag cgc cga ccc caa ggt tta ccc aat   144
Arg Asn Gly Ala Arg Pro Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn
        35   40   45

aat act gcg tct tgg ttc aca gct ctc act cag cat ggc aag gag gaa   192
Asn Thr Ala Ser Trp Phe Thr Ala Leu Thr Gln His Gly Lys Glu Glu
    50   55   60

ctt aga ttc cct cga ggc cag ggc gtt cca atc aac acc aat agt ggt   240
Leu Arg Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Gly
65   70   75   80

cca gat gac caa att ggc tac tac cga aga gct acc cga cga gtt cgt   288
Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Val Arg
                85   90   95

1

```
ggt ggt gac ggc aaa atg aaa gag ctc agc ccc aga tgg tac ttc tat        336
Gly Gly Asp Gly Lys Met Lys Glu Leu Ser Pro Arg Trp Tyr Phe Tyr
            100             105                 110

tac cta gga act ggc cca gaa gct tca ctt ccc tac ggc gct aac aaa        384
Tyr Leu Gly Thr Gly Pro Glu Ala Ser Leu Pro Tyr Gly Ala Asn Lys
            115             120                 125

gaa ggc atc gta tgg gtt gca act gag gga gcc ttg aat aca ccc aaa        432
Glu Gly Ile Val Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys
        130             135                 140

gac cac att ggc acc cgc aat cct aat aac aat gct gcc acc gtg cta        480
Asp His Ile Gly Thr Arg Asn Pro Asn Asn Asn Ala Ala Thr Val Leu
145                 150                 155                 160

caa ctt cct caa gga aca aca ttg cca aaa ggc ttc tac gca gag gga        528
Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly
                165                 170                 175

agc aga ggc ggc agt caa gcc tct tct cgc tcc tca tca cgt agt cgc        576
Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg
                180                 185                 190

ggt aat tca aga aat tca act cct ggc agc agt agg gga aat tct cct        624
Gly Asn Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Asn Ser Pro
                195                 200                 205

gct cga atg gct agc gga ggt ggt gaa act gcc ctc gcg cta ttg ctg        672
Ala Arg Met Ala Ser Gly Gly Gly Glu Thr Ala Leu Ala Leu Leu Leu
        210                 215                 220

cta gac aga ttg aac cag ctt gag agc aaa gtt tct ggt aaa ggc caa        720
Leu Asp Arg Leu Asn Gln Leu Glu Ser Lys Val Ser Gly Lys Gly Gln
225                 230                 235                 240

caa caa caa ggc caa act gtc act aag aaa tct gct gct gag gca tct        768
Gln Gln Gln Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser
                245                 250                 255

aaa aag cct cgc caa aaa cgt act gcc aca aaa cag tac aac gtc act        816
Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr
                260                 265                 270

caa gca ttt ggg aga cgt ggt cca gaa caa acc caa gga aat ttc ggg        864
Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly
                275                 280                 285

gac caa gac cta atc aga caa gga act gat tac aaa cat tgg ccg caa        912
Asp Gln Asp Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln
            290                 295                 300

att gca caa ttt gct cca agt gcc tct gca ttc ttt gga atg tca cgc        960
Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg
```

2

```
att ggc atg gaa gtc aca cct tcg gga aca tgg ctg act tat cat gga      1008
Ile Gly Met Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr His Gly
                325                  330                  335

gcc att aaa ttg gat gac aaa gat cca caa ttc aaa gac aac gtc ata      1056
Ala Ile Lys Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile
                340                  345                  350

ctg ctg aac aag cac att gac gca tac aaa aca ttc cca cca aca gag      1104
Leu Leu Asn Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu
                355                  360                  365

cct aaa aag gac aaa aag aaa aag act gat gaa gct cag cct ttg ccg      1152
Pro Lys Lys Asp Lys Lys Lys Lys Thr Asp Glu Ala Gln Pro Leu Pro
            370                  375                  380

cag aga caa aag aag cag ccc act gtg act ctt ctt cct gcg gct gac      1200
Gln Arg Gln Lys Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp
385                  390                  395                  400

atg gat gat ttc tcc aga caa ctt caa aat tcc atg agt gga gct tct      1248
Met Asp Asp Phe Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser
                405                  410                  415

gct gat tca act cag gca taa                                          1269
Ala Asp Ser Thr Gln Ala
                420
```

<210> 2
<211> 422
<212> PRT
<213> SARS coronavirus

<400> 2

Met Ser Asp Asn Gly Pro Gln Ser Asn Gln Arg Ser Ala Pro Arg Ile
1   5   10   15

Thr Phe Gly Gly Pro Thr Asp Ser Thr Asp Asn Asn Gln Asn Gly Gly
            20   25   30

Arg Asn Gly Ala Arg Pro Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn
            35   40   45

Asn Thr Ala Ser Trp Phe Thr Ala Leu Thr Gln His Gly Lys Glu Glu
        50   55   60

Leu Arg Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Gly

Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Val Arg
            85   90   95

Gly Gly Asp Gly Lys Met Lys Glu Leu Ser Pro Arg Trp Tyr Phe Tyr
            100   105   110

Tyr Leu Gly Thr Gly Pro Glu Ala Ser Leu Pro Tyr Gly Ala Asn Lys
            115   120   125

Glu Gly Ile Val Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys
            130   135   140

Asp His Ile Gly Thr Arg Asn Pro Asn Asn Asn Ala Ala Thr Val Leu
145   150   155   160

Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly
            165   170   175

Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg
            180   185   190

Gly Asn Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Asn Ser Pro
            195   200   205

Ala Arg Met Ala Ser Gly Gly Gly Glu Thr Ala Leu Ala Leu Leu Leu
            210   215   220

Leu Asp Arg Leu Asn Gln Leu Glu Ser Lys Val Ser Gly Lys Gly Gln
225   230   235   240

Gln Gln Gln Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser
            245   250   255

Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr
            260   265   270

Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly
            275   280   285

```
Asp Gln Asp Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln
    290                 295             300

Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg
305             310             315                 320

Ile Gly Met Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr His Gly
            325             330             335

Ala Ile Lys Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile
        340             345             350

Leu Leu Asn Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu
        355             360             365

Pro Lys Lys Asp Lys Lys Lys Lys Thr Asp Glu Ala Gln Pro Leu Pro
    370             375             380

Gln Arg Gln Lys Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp
385             390             395                 400

Met Asp Asp Phe Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser
            405             410             415

Ala Asp Ser Thr Gln Ala
            420


<210>   3
<211>   3768
<212>   DNA
<213>   SARS coronavirus


<220>
<221>   CDS
<222>   (1)..(3768)

<400>   3
atg ttt att ttc tta tta ttt ctt act ctc act agt ggt agt gac ctt      48
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1               5                   10                  15

gac cgg tgc acc act ttt gat gat gtt caa gct cct aat tac act caa      96
Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
            20                  25                  30
```

```
cat act tca tct atg agg ggg gtt tac tat cct gat gaa att ttt aga      144
His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
        35                      40                      45

tca gac act ctt tat tta act cag gat tta ttt ctt cca ttt tat tct      192
Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
        50                      55                      60

aat gtt aca ggg ttt cat act att aat cat acg ttt ggc aac cct gtc      240
Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65                      70                      75                      80

ata cct ttt aag gat ggt att tat ttt gct gcc aca gag aaa tca aat      288
Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
                85                      90                      95

gtt gtc cgt ggt tgg gtt ttt ggt tct acc atg aac aac aag tca cag      336
Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
            100                     105                     110

tcg gtg att att att aac aat tct act aat gtt gtt ata cga gca tgt      384
Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
            115                     120                     125

aac ttt gaa ttg tgt gac aac cct ttc ttt gct gtt tct aaa ccc atg      432
Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
        130                     135                     140

ggt aca cag aca cat act atg ata ttc gat aat gca ttt aat tgc act      480
Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145                     150                     155                     160

ttc gag tac ata tct gat gcc ttt tcg ctt gat gtt tca gaa aag tca      528
Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
                165                     170                     175

ggt aat ttt aaa cac tta cga gag ttt gtg ttt aaa aat aaa gat ggg      576
Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
            180                     185                     190

ttt ctc tat gtt tat aag ggc tat caa cct ata gat gta gtt cgt gat      624
Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
            195                     200                     205

cta cct tct ggt ttt aac act ttg aaa cct att ttt aag ttg cct ctt      672
Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
        210                     215                     220

ggt att aac att aca aat ttt aga gcc att ctt aca gcc ttt tca cct      720
Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225                     230                     235                     240

gct caa gac att tgg ggc acg tca gct gca gcc tat ttt gtt ggc tat      768
Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
                245                     250                     255
```

6

```
tta aag cca act aca ttt atg ctc aag tat gat gaa aat ggt aca atc        816
Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
            260             265             270

aca gat gct gtt gat tgt tct caa aat cca ctt gct gaa ctc aaa tgc        864
Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
            275             280             285

tct gtt aag agc ttt gag att gac aaa gga att tac cag acc tct aat        912
Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
            290             295             300

ttc agg gtt gtt ccc tca gga gat gtt gtg aga ttc cct aat att aca        960
Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305             310             315             320

aac ttg tgt cct ttt gga gag gtt ttt aat gct act aaa ttc cct tct       1008
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
            325             330             335

gtc tat gca tgg gag aga aaa aaa att tct aat tgt gtt gct gat tac       1056
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
            340             345             350

tct gtg ctc tac aac tca aca ttt ttt tca acc ttt aag tgc tat ggc       1104
Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
            355             360             365

gtt tct gcc act aag ttg aat gat ctt tgc ttc tcc aat gtc tat gca       1152
Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
            370             375             380

gat tct ttt gta gtc aag gga gat gat gta aga caa ata gcg cca gga       1200
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385             390             395             400

caa act ggt gtt att gct gat tat aat tat aaa ttg cca gat gat ttc       1248
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
            405             410             415

atg ggt tgt gtc ctt gct tgg aat act agg aac att gat gct act tca       1296
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
            420             425             430

act ggt aat tat aat tat aaa tat agg tat ctt aga cat ggc aag ctt       1344
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
            435             440             445

agg ccc ttt gag aga gac ata tct aat gtg cct ttc tcc cct gat ggc       1392
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
            450             455             460

aaa cct tgc acc cca cct gct ctt aat tgt tat tgg cca tta aat gat       1440
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
```

7

tat ggt ttt tac acc act act ggc att ggc tac caa cct tac aga gtt   1488
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
                485   490   495

gta gta ctt tct ttt gaa ctt tta aat gca ccg gcc acg gtt tgt gga   1536
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
                500   505   510

cca aaa tta tcc act gac ctt att aag aac cag tgt gtc aat ttt aat   1584
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
                515   520   525

ttt aat gga ctc act ggt act ggt gtg tta act cct tct tca aag aga   1632
Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
                530   535   540

ttt caa cca ttt caa caa ttt ggc cgt gat gtt tct gat ttc act gat   1680
Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
545   550   555   560

tcc gtt cga gat cct aaa aca tct gaa ata tta gac att tca cct tgc   1728
Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
                565   570   575

tct ttt ggg ggt gta agt gta att aca cct gga aca aat gct tca tct   1776
Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
                580   585   590

gaa gtt gct gtt cta tat caa gat gtt aac tgc act gat gtt tct aca   1824
Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
                595   600   605

gca att cat gca gat caa ctc aca cca gct tgg cgc ata tat tct act   1872
Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
                610   615   620

gga aac aat gta ttc cag act caa gca ggc tgt ctt ata gga gct gag   1920
Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
625   630   635   640

cat gtc gac act tct tat gag tgc gac att cct att gga gct ggc att   1968
His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
                645   650   655

tgt gct agt tac cat aca gtt tct tta tta cgt agt act agc caa aaa   2016
Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
                660   665   670

tct att gtg gct tat act atg tct tta ggt gct gat agt tca att gct   2064
Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
                675   680   685

tac tct aat aac acc att gct ata cct act aac ttt tca att agc att   2112

8

```
Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
    690                 695             700

act aca gaa gta atg cct gtt tct atg gct aaa acc tcc gta gat tgt      2160
Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
705             710             715                 720

aat atg tac atc tgc gga gat tct act gaa tgt gct aat ttg ctt ctc      2208
Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
                725             730             735

caa tat ggt agc ttt tgc aca caa cta aat cgt gca ctc tca ggt att      2256
Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
            740             745             750

gct gct gaa cag gat cgc aac aca cgt gaa gtg ttc gct caa gtt aaa      2304
Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
            755             760             765

caa atg tac aaa acc cca act ttg aaa tat ttt ggt ggt ttt aat ttt      2352
Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
        770             775             780

tca caa ata tta cct gac cct cta aag cca act aag agg tct ttt att      2400
Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785             790             795                 800

gag gac ttg ctc ttt aat aag gtg aca ctc gct gat gct ggc ttc atg      2448
Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
            805             810             815

aag caa tat ggc gaa tgc cta ggt gat att aat gct aga gat ctc att      2496
Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
            820             825             830

tgt gcg cag aag ttc aat gga ctt aca gtg ttg cca cct ctg ctc act      2544
Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
            835             840             845

gat gat atg att gct gcc tac act gct gct cta gtt agt ggt act gcc      2592
Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
            850             855             860

act gct gga tgg aca ttt ggt gct ggc gct gct ctt caa ata cct ttt      2640
Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
865             870             875                 880

gct atg caa atg gca tat agg ttc aat ggc att gga gtt acc caa aat      2688
Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
            885             890             895

gtt ctc tat gag aac caa aaa caa atc gcc aac caa ttt aac aag gcg      2736
Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
        900             905             910
```

9

```
att agt caa att caa gaa tca ctt aca aca aca tca act gca ttg ggc          2784
Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
        915                 920                 925

aag ctg caa gac gtt gtt aac cag aat gct caa gca tta aac aca ctt          2832
Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
    930                 935                 940

gtt aaa caa ctt agc tct aat ttt ggt gca att tca agt gtg cta aat          2880
Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
945             950         .     955                 960

gat atc ctt tcg cga ctt gat aaa gtc gag gcg gag gta caa att gac          2928
Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
                965                 970                 975

agg tta att aca ggc aga ctt caa agc ctt caa acc tat gta aca caa          2976
Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
            980                 985                 990

caa cta atc agg gct gct gaa atc  agg gct tct gct aat  ctt gct gct        3024
Gln Leu Ile Arg Ala Ala Glu Ile  Arg Ala Ser Ala Asn  Leu Ala Ala
        995                 1000                1005

act aaa  atg tct gag tgt gtt  ctt gga caa tca aaa  aga gtt gac          3069
Thr Lys  Met Ser Glu Cys Val  Leu Gly Gln Ser Lys  Arg Val Asp
    1010             1015                1020

ttt tgt  gga aag ggc tac cac  ctt atg tcc ttc cca  caa gca gcc          3114
Phe Cys  Gly Lys Gly Tyr His  Leu Met Ser Phe Pro  Gln Ala Ala
    1025             1030                1035

ccg cat  ggt gtt gtc ttc cta  cat gtc acg tat gtg  cca tcc cag          3159
Pro His  Gly Val Val Phe Leu  His Val Thr Tyr Val  Pro Ser Gln
    1040             1045                1050

gag agg  aac ttc acc aca gcg  cca gca att tgt cat  gaa ggc aaa          3204
Glu Arg  Asn Phe Thr Thr Ala  Pro Ala Ile Cys His  Glu Gly Lys
    1055             1060                1065

gca tac  ttc cct cgt gaa ggt  gtt ttt gtg ttt aat  ggc act tct          3249
Ala Tyr  Phe Pro Arg Glu Gly  Val Phe Val Phe Asn  Gly Thr Ser
    1070             1075                1080

tgg ttt  att aca cag agg aac  ttc ttt tct cca caa  ata att act          3294
Trp Phe  Ile Thr Gln Arg Asn  Phe Phe Ser Pro Gln  Ile Ile Thr
    1085             1090                1095

aca gac  aat aca ttt gtc tca  gga aat tgt gat gtc  gtt att ggc          3339
Thr Asp  Asn Thr Phe Val Ser  Gly Asn Cys Asp Val  Val Ile Gly
    1100             1105                1110

atc att  aac aac aca gtt tat  gat cct ctg caa cct  gag ctt gac          3384
Ile Ile  Asn Asn Thr Val Tyr  Asp Pro Leu Gln Pro  Glu Leu Asp
    1115             1120                1125
```

10

```
tca ttc aaa gaa gag ctg gac  aag tac ttc aaa aat  cat aca tca       3429
Ser Phe Lys Glu Glu Leu Asp  Lys Tyr Phe Lys Asn  His Thr Ser
    1130             1135              1140

cca gat gtt gat ctt ggc gac  att tca ggc att aac  gct tct gtc       3474
Pro Asp Val Asp Leu Gly Asp  Ile Ser Gly Ile Asn  Ala Ser Val
    1145             1150              1155

gtc aac att caa aaa gaa att  gac cgc ctc aat gag  gtc gct aaa       3519
Val Asn Ile Gln Lys Glu Ile  Asp Arg Leu Asn Glu  Val Ala Lys
    1160             1165              1170

aat tta aat gaa tca ctc att  gac ctt caa gaa ttg  gga aaa tat       3564
Asn Leu Asn Glu Ser Leu Ile  Asp Leu Gln Glu Leu  Gly Lys Tyr
    1175             1180              1185

gag caa tat att aaa tgg cct  tgg tat gtt tgg ctc  ggc ttc att       3609
Glu Gln Tyr Ile Lys Trp Pro  Trp Tyr Val Trp Leu  Gly Phe Ile
    1190             1195              1200

gct gga cta att gcc atc gtc  atg gtt aca atc ttg  ctt tgt tgc       3654
Ala Gly Leu Ile Ala Ile Val  Met Val Thr Ile Leu  Leu Cys Cys
    1205             1210              1215

atg act agt tgt tgc agt tgc  ctc aag ggt gca tgc  tct tgt ggt       3699
Met Thr Ser Cys Cys Ser Cys  Leu Lys Gly Ala Cys  Ser Cys Gly
    1220             1225              1230

tct tgc tgc aag ttt gat gag  gat gac tct gag cca  gtt ctc aag       3744
Ser Cys Cys Lys Phe Asp Glu  Asp Asp Ser Glu Pro  Val Leu Lys
    1235             1240              1245

ggt gtc aaa tta cat tac aca  taa                                    3768
Gly Val Lys Leu His Tyr Thr
    1250             1255


<210>   4
<211>   1255
<212>   PRT
<213>   SARS coronavirus

<400>   4

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1               5                   10                  15


Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
            20                  25                  30


His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
            35                  40                  45
```

11

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
    50   55   60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65   70   75   80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
            85   90   95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
        100   105   110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
        115   120   125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
    130   135   140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145   150   155   160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
            165   170   175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
        180   185   190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
        195   200   205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
    210   215   220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225   230   235   240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
            245   250   255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile

12

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
    275   280   285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
    290   295   300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305   310   315   320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
            325   330   335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
            340   345   350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
    355   360   365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
    370   375   380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
385   390   395   400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
            405   410   415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
            420   425   430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
    435   440   445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
    450   455   460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
465   470   475   480

13

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
                      485   490   495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
              500   505   510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
          515   520   525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
      530   .   535   540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
545   550   555   560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
              565   570   575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
          580   585   590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
          595   600   605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
      610   615   620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
625   630   635   640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
              645   650   655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
          660   665   670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
          675   680   685

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
      690   695   700

14

```
Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
705             710             715             720


Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
            725             730             735


Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
        740             745             750


Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
        755             760             765


Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
    770             775             780


Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785             790             795             800


Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
            805             810             815


Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
        820             825             830


Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
        835             840             845


Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
    850             855             860


Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
865             870             875             880


Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
            885             890             895


Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
        900             905             910


Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
        915             920             925
```

15

Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
    930   935   940

Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
945   950   955   960

Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
                965   970   975

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
                980   985   990

Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
        995   1000   1005

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp
    1010   1015   1020

Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala
    1025   1030   1035

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln
    1040   1045   1050

Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys
    1055   1060   1065

Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser
    1070   1075   1080

Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr
    1085   1090   1095

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
    1100   1105   1110

Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
    1115   1120   1125

Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser

```
                1130                      1135                      1140


Pro Asp  Val Asp Leu Gly Asp  Ile Ser Gly Ile Asn  Ala Ser Val
    1145                      1150                      1155


Val Asn  Ile Gln Lys Glu Ile  Asp Arg Leu Asn Glu  Val Ala Lys
    1160                      1165.                     1170


Asn Leu  Asn Glu Ser Leu Ile  Asp Leu Gln Glu Leu  Gly Lys Tyr
    1175                      1180                      1185


Glu Gln  Tyr Ile Lys Trp Pro  Trp Tyr Val Trp Leu  Gly Phe Ile
    1190                      1195                      1200


Ala Gly  Leu Ile Ala Ile Val  Met Val Thr Ile Leu  Leu Cys Cys
    1205                      1210                      1215


Met Thr  Ser Cys Cys Ser Cys  Leu Lys Gly Ala Cys  Ser Cys Gly
    1220                      1225                      1230


Ser Cys  Cys Lys Phe Asp Glu  Asp Asp Ser Glu Pro  Val Leu Lys
    1235                      1240                      1245


Gly Val  Lys Leu His Tyr Thr
    1250                      1255



<210>   5
<211>   588
<212>   DNA
<213>   SARS coronavirus


<220>
<221>   CDS
<222>   (1)..(588)

<400>   5
ttg aac cag ctt gag agc aaa gtt tct ggt aaa ggc caa caa caa caa        48
Leu Asn Gln Leu Glu Ser Lys Val Ser Gly Lys Gly Gln Gln Gln Gln
1               5                   10                  15


ggc caa act gtc act aag aaa tct gct gct gag gca tct aaa aag cct        96
Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser Lys Lys Pro
            20                  25                  30


cgc caa aaa cgt act gcc aca aaa cag tac aac gtc act caa gca ttt       144
```

17

```
Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr Gln Ala Phe
        35              40              45

ggg aga cgt ggt cca gaa caa acc caa gga aat ttc ggg gac caa gac        192
Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly Asp Gln Asp
    50              55              60

cta atc aga caa gga act gat tac aaa cat tgg ccg caa att gca caa        240
Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln Ile Ala Gln
65              70              75              80

ttt gct cca agt gcc tct gca ttc ttt gga atg tca cgc att ggc atg        288
Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg Ile Gly Met
            85              90              95

gaa gtc aca cct tcg gga aca tgg ctg act tat cat gga gcc att aaa        336
Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr His Gly Ala Ile Lys
            100             105             110

ttg gat gac aaa gat cca caa ttc aaa gac aac gtc ata ctg ctg aac        384
Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile Leu Leu Asn
            115             120             125

aag cac att gac gca tac aaa aca ttc cca cca aca gag cct aaa aag        432
Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu Pro Lys Lys
    130             135             140

gac aaa aag aaa aag act gat gaa gct cag cct ttg ccg cag aga caa        480
Asp Lys Lys Lys Lys Thr Asp Glu Ala Gln Pro Leu Pro Gln Arg Gln
145             150             155             160

aag aag cag ccc act gtg act ctt ctt cct gcg gct gac atg gat gat        528
Lys Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp Met Asp Asp
            165             170             175

ttc tcc aga caa ctt caa aat tcc atg agt gga gct tct gct gat tca        576
Phe Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser Ala Asp Ser
            180             185             190

act cag gca taa                                                        588
Thr Gln Ala
        195


<210>   6
<211>   195
<212>   PRT
<213>   SARS coronavirus

<400>   6

Leu Asn Gln Leu Glu Ser Lys Val Ser Gly Lys Gly Gln Gln Gln Gln
1               5               10              15
```

18

Gly Gln Thr Val Thr Lys Lys Ser Ala Ala Glu Ala Ser Lys Lys Pro
            20   25   30

Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr Gln Ala Phe
            35   40   45

Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly Asp Gln Asp
            50   55   60

Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln Ile Ala Gln
65   70   75   80

Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg Ile Gly Met
            85   90   95

Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr His Gly Ala Ile Lys
            100   105   110

Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile Leu Leu Asn
            115   120   125

Lys His Ile Asp Ala Tyr Lys Thr Phe Pro Pro Thr Glu Pro Lys Lys
    130   135   140

Asp Lys Lys Lys Lys Thr Asp Glu Ala Gln Pro Leu Pro Gln Arg Gln
145   150   155   160

Lys Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp Met Asp Asp
            165   170   175

Phe Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser Ala Asp Ser
            180   185   190

Thr Gln Ala
        195

<210> 7
<211> 684
<212> DNA
<213> SARS coronavirus

<220>

19

<221>   CDS
<222>   (1)..(684)

<400>   7

```
agg tat ctt aga cat ggc aag ctt agg ccc ttt gag aga gac ata tct          48
Arg Tyr Leu Arg His Gly Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser
1               5                   10                  15

aat gtg cct ttc tcc cct gat ggc aaa cct tgc acc cca cct gct ctt          96
Asn Val Pro Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu
            20                  25                  30

aat tgt tat tgg cca tta aat gat tat ggt ttt tac acc act act ggc         144
Asn Cys Tyr Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly
        35                  40                  45

att ggc tac caa cct tac aga gtt gta gta ctt tct ttt gaa ctt tta         192
Ile Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu
    50                  55                  60

aat gca ccg gcc acg gtt tgt gga cca aaa tta tcc act gac ctt att         240
Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile
65                  70                  75                  80

aag aac cag tgt gtc aat ttt aat ttt aat gga ctc act ggt act ggt         288
Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly
                85                  90                  95

gtg tta act cct tct tca aag aga ttt caa cca ttt caa caa ttt ggc         336
Val Leu Thr Pro Ser Ser Lys Arg Phe Gln Pro Phe Gln Gln Phe Gly
            100                 105                 110

cgt gat gtt tct gat ttc act gat tcc gtt cga gat cct aaa aca tct         384
Arg Asp Val Ser Asp Phe Thr Asp Ser Val Arg Asp Pro Lys Thr Ser
        115                 120                 125

gaa ata tta gac att tca cct tgc tct ttt ggg ggt gta agt gta att         432
Glu Ile Leu Asp Ile Ser Pro Cys Ser Phe Gly Gly Val Ser Val Ile
    130                 135                 140

aca cct gga aca aat gct tca tct gaa gtt gct gtt cta tat caa gat         480
Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val Leu Tyr Gln Asp
145                 150                 155                 160

gtt aac tgc act gat gtt tct aca gca att cat gca gat caa ctc aca         528
Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr
                165                 170                 175

cca gct tgg cgc ata tat tct act gga aac aat gta ttc cag act caa         576
Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln
            180                 185                 190

gca ggc tgt ctt ata gga gct gag cat gtc gac act tct tat gag tgc         624
Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys
        195                 200                 205
```

20

```
gac att cct att gga gct ggc att tgt gct agt tac cat aca gtt tct       672
Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser
    210             215             220

tta tta cgt agt                                                       684
Leu Leu Arg Ser
225


<210>   8
<211>   228
<212>   PRT
<213>   SARS coronavirus

<400>   8

Arg Tyr Leu Arg His Gly Lys Leu Arg Pro Phe Glu Arg Asp Ile Ser
1               5               10              15


Asn Val Pro Phe Ser Pro Asp Gly Lys Pro Cys Thr Pro Pro Ala Leu
            20              25              30


Asn Cys Tyr Trp Pro Leu Asn Asp Tyr Gly Phe Tyr Thr Thr Thr Gly
        35              40              45


Ile Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu
    50              55              60


Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser Thr Asp Leu Ile
65              70              75              80


Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu Thr Gly Thr Gly
                85              90              95


Val Leu Thr Pro Ser Ser Lys Arg Phe Gln Pro Phe Gln Gln Phe Gly
            100             105             110


Arg Asp Val Ser Asp Phe Thr Asp Ser Val Arg Asp Pro Lys Thr Ser
        115             120             125


Glu Ile Leu Asp Ile Ser Pro Cys Ser Phe Gly Gly Val Ser Val Ile
    130             135             140


Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val Leu Tyr Gln Asp
145             150             155             160


                                     21
```

Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala Asp Gln Leu Thr
            165   170   175

Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val Phe Gln Thr Gln
            180   185   190

Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr Ser Tyr Glu Cys
            195   200   205

Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr His Thr Val Ser
    210   215   220

Leu Leu Arg Ser
225

<210>   9
<211>   29711
<212>   DNA
<213>   SARS coronavirus

<400>   9
tacccaggaa aagccaacca acctcgatct cttgtagatc tgttctctaa acgaacttta   60

aaatctgtgt agctgtcgct cggctgcatg cctagtgcac ctacgcagta taaacaataa   120

taaattttac tgtcgttgac aagaaacgag taactcgtcc ctcttctgca gactgcttac   180

ggtttcgtcc gtgttgcagt cgatcatcag catacctagg tttcgtccgg gtgtgaccga   240

aaggtaagat ggagagcctt gttcttggtg tcaacgagaa aacacacgtc caactcagtt   300

tgcctgtcct tcaggttaga gacgtgctag tgcgtggctt cggggactct gtggaagagg   360

ccctatcgga ggcacgtgaa cacctcaaaa atggcacttg tggtctagta gagctggaaa   420

aaggcgtact gccccagctt gaacagccct atgtgttcat taaacgttct gatgccttaa   480

gcaccaatca cggccacaag gtcgttgagc tggttgcaga aatggacggc attcagtacg   540

gtcgtagcgg tataacactg ggagtactcg tgccacatgt gggcgaaacc ccaattgcat   600

accgcaatgt tcttcttcgt aagaacggta ataagggagc cggtggtcat agctatggca   660

tcgatctaaa gtcttatgac ttaggtgacg agcttggcac tgatcccatt gaagattatg   720

aacaaaactg gaacactaag catggcagtg gtgcactccg tgaactcact cgtgagctca   780

atggaggtgc agtcactcgc tatgtcgaca acaatttctg tggcccagat gggtaccctc   840

22

```
ttgattgcat caaagatttt ctcgcacgcg cgggcaagtc aatgtgcact ctttccgaac    900

aacttgatta catcgagtcg aagagaggtg tctactgctg ccgtgaccat gagcatgaaa    960

ttgcctggtt cactgagcgc tctgataaga gctacgagca ccagacaccc ttcgaaatta   1020

agagtgccaa gaaatttgac actttcaaag gggaatgccc aaagtttgtg tttcctctta   1080

actcaaaagt caaagtcatt caaccacgtg ttgaaaagaa aaagactgag ggtttcatgg   1140

ggcgtatacg ctctgtgtac cctgttgcat ctccacagga gtgtaacaat atgcacttgt   1200

ctaccttgat gaaatgtaat cattgcgatg aagtttcatg gcagacgtgc gactttctga   1260

aagccacttg tgaacattgt ggcactgaaa atttagttat tgaaggacct actacatgtg   1320

ggtacctacc tactaatgct gtagtgaaaa tgccatgtcc tgcctgtcaa gacccagaga   1380

ttggacctga gcatagtgtt gcagattatc acaaccactc aaacattgaa actcgactcc   1440

gcaagggagg taggactaga tgttttggag gctgtgtgtt tgcctatgtt ggctgctata   1500

ataagcgtgc ctactgggtt cctcgtgcta gtgctgatat tggctcaggc catactggca   1560

ttactggtga caatgtggag accttgaatg aggatctcct tgagatactg agtcgtgaac   1620

gtgttaacat taacattgtt ggcgattttc atttgaatga agaggttgcc atcattttgg   1680

catctttctc tgcttctaca agtgccttta ttgacactat aaagagtctt gattacaagt   1740

ctttcaaaac cattgttgag tcctgcggta actataaagt taccaaggga aagcccgtaa   1800

aaggtgcttg gaacattgga caacagagat cagtttttaac accactgtgt ggttttccct   1860

cacaggctgc tggtgttatc agatcaattt ttgcgcgcac acttgatgca gcaaaccact   1920

caattcctga tttgcaaaga gcagctgtca ccatacttga tggtatttct gaacagtcat   1980

tacgtcttgt cgacgccatg gtttatactt cagacctgct caccaacagt gtcattatta   2040

tggcatatgt aactggtggt cttgtacaac agacttctca gtggttgtct aatcttttgg   2100

gcactactgt tgaaaaactc aggcctatct ttgaatggat tgaggcgaaa cttagtgcag   2160

gagttgaatt tctcaaggat gcttgggaga ttctcaaatt tctcattaca ggtgttttg    2220

acatcgtcaa gggtcaaata caggttgctt cagataacat caaggattgt gtaaaatgct   2280

tcattgatgt tgttaacaag gcactcgaaa tgtgcattga tcaagtcact atcgctggcg   2340

caaagttgcg atcactcaac ttaggtgaag tcttcatcgc tcaaagcaag ggactttacc   2400

gtcagtgtat acgtggcaag gagcagctgc aactactcat gcctcttaag gcaccaaaag   2460
```

23

```
aagtaacctt tcttgaaggt gattcacatg acacagtact tacctctgag gaggttgttc    2520

tcaagaacgg tgaactcgaa gcactcgaga cgcccgttga tagcttcaca aatggagcta    2580

tcgttggcac accagtctgt gtaaatggcc tcatgctctt agagattaag gacaaagaac    2640

aatactgcgc attgtctcct ggtttactgg ctacaaacaa tgtctttcgc ttaaaagggg    2700

gtgcaccaat taaaggtgta acctttggag aagatactgt ttgggaagtt caaggttaca    2760

agaatgtgag aatcacattt gagcttgatg aacgtgttga caaagtgctt aatgaaaagt    2820

gctctgtcta cactgttgaa tccggtaccg aagttactga gtttgcatgt gttgtagcag    2880

aggctgttgt gaagacttta caaccagttt ctgatctcct taccaacatg ggtattgatc    2940

ttgatgagtg gagtgtagct acattctact tatttgatga tgctggtgaa gaaaactttt    3000

catcacgtat gtattgttcc ttttaccctc cagatgagga agaagaggac gatgcagagt    3060

gtgaggaaga agaaattgat gaaacctgtg aacatgagta cggtacagag gatgattatc    3120

aaggtctccc tctggaattt ggtgcctcag ctgaaacagt tcgagttgag gaagaagaag    3180

aggaagactg gctggatgat actactgagc aatcagagat tgagccagaa ccagaaccta    3240

cacctgaaga accagttaat cagtttactg gttatttaaa acttactgac aatgttgcca    3300

ttaaatgtgt tgacatcgtt aaggaggcac aaagtgctaa tcctatggtg attgtaaatg    3360

ctgctaacat acacctgaaa catggtggtg gtgtagcagg tgcactcaac aaggcaacca    3420

atggtgccat gcaaaaggag agtgatgatt acattaagct aaatggccct cttacagtag    3480

gagggtcttg tttgctttct ggacataatc ttgctaagaa gtgtctgcat gttgttggac    3540

ctaacctaaa tgcaggtgag gacatccagc ttcttaaggc agcatatgaa aatttcaatt    3600

cacaggacat cttacttgca ccattgttgt cagcaggcat atttggtgct aaaccacttc    3660

agtctttaca agtgtgcgtg cagacggttc gtacacaggt ttatattgca gtcaatgaca    3720

aagctcttta tgagcaggtt gtcatggatt atcttgataa cctgaagcct agagtggaag    3780

cacctaaaca agaggagcca ccaaacacag aagattccaa aactgaggag aaatctgtcg    3840

tacagaagcc tgtcgatgtg aagccaaaaa ttaaggcctg cattgatgag gttaccacaa    3900

cactggaaga aactaagttt cttaccaata agttactctt gtttgctgat atcaatggta    3960

agctttacca tgattctcag aacatgctta gaggtgaaga tatgtctttc cttgagaagg    4020

atgcacctta catggtaggt gatgttatca ctagtggtga tatcacttgt gttgtaatac    4080

cctccaaaaa ggctggtggc actactgaga tgctctcaag agctttgaag aaagtgccag    4140
```

24

```
ttgatgagta tataaccacg taccctggac aaggatgtgc tggttataca cttgaggaag    4200

ctaagactgc tcttaagaaa tgcaaatctg cattttatgt actaccttca gaagcaccta    4260

atgctaagga agagattcta ggaactgtat cctggaattt gagagaaatg cttgctcatg    4320

ctgaagagac aagaaaatta atgcctatat gcatggatgt tagagccata atggcaacca    4380

tccaacgtaa gtataaagga attaaaattc aagagggcat cgttgactat ggtgtccgat    4440

tcttctttta tactagtaaa gagcctgtag cttctattat tacgaagctg aactctctaa    4500

atgagccgct tgtcacaatg ccaattggtt atgtgacaca tggttttaat cttgaagagg    4560

ctgcgcgctg tatgcgttct cttaaagctc ctgccgtagt gtcagtatca tcaccagatg    4620

ctgttactac atataatgga tacctcactt cgtcatcaaa gacatctgag gagcactttg    4680

tagaaacagt ttctttggct ggctcttaca gagattggtc ctattcagga cagcgtacag    4740

agttaggtgt tgaatttctt aagcgtggtg acaaaattgt gtaccacact ctggagagcc    4800

ccgtcgagtt tcatcttgac ggtgaggttc tttcacttga caaactaaag agtctcttat    4860

ccctgcggga ggttaagact ataaaagtgt tcacaactgt ggacaacact aatctccaca    4920

cacagcttgt ggatatgtct atgacatatg gacagcagtt tggtccaaca tacttggatg    4980

gtgctgatgt tacaaaaatt aaacctcatg taaatcatga gggtaagact ttctttgtac    5040

tacctagtga tgacacacta cgtagtgaag ctttcgagta ctaccatact cttgatgaga    5100

gttttcttgg taggtacatg tctgctttaa accacacaaa gaaatggaaa tttcctcaag    5160

ttggtggttt aacttcaatt aaatgggctg ataacaattg ttatttgtct agtgttttat    5220

tagcacttca acagcttgaa gtcaaattca atgcaccagc acttcaagag gcttattata    5280

gagcccgtgc tggtgatgct gctaactttt gtgcactcat actcgcttac agtaataaaa    5340

ctgttggcga gcttggtgat gtcagagaaa ctatgaccca tcttctacag catgctaatt    5400

tggaatctgc aaagcgagtt cttaatgtgg tgtgtaaaca ttgtggtcag aaaactacta    5460

ccttaacggg tgtagaagct gtgatgtata tgggtactct atcttatgat aatcttaaga    5520

caggtgtttc cattccatgt gtgtgtggtc gtgatgctac acaatatcta gtacaacaag    5580

agtcttcttt tgttatgatg tctgcaccac ctgctgagta taaattacag caaggtacat    5640

tcttatgtgc gaatgagtac actggtaact atcagtgtgg tcattacact catataactg    5700

ctaaggagac cctctatcgt attgacggag ctcaccttac aaagatgtca gagtacaaag    5760
```

25

```
gaccagtgac tgatgttttc tacaaggaaa catcttacac tacaaccatc aagcctgtgt    5820

cgtataaact cgatggagtt acttacacag agattgaacc aaaattggat gggtattata    5880

aaaaggataa tgcttactat acagagcagc ctatagacct tgtaccaact caaccattac    5940

caaatgcgag ttttgataat ttcaaactca catgttctaa cacaaaattt gctgatgatt    6000

taaatcaaat gacaggcttc acaaagccag cttcacgaga gctatctgtc acattcttcc    6060

cagacttgaa tggcgatgta gtggctattg actatagaca ctattcagcg agtttcaaga    6120

aaggtgctaa attactgcat aagccaattg tttggcacat taaccaggct acaaccaaga    6180

caacgttcaa accaaacact tggtgtttac gttgtctttg gagtacaaag ccagtagata    6240

cttcaaattc atttgaagtt ctggcagtag aagacacaca aggaatggac aatcttgctt    6300

gtgaaagtca acaacccacc tctgaagaag tagtggaaaa tcctaccata cagaaggaag    6360

tcatagagtg tgacgtgaaa actaccgaag ttgtaggcaa tgtcatactt aaaccatcag    6420

atgaaggtgt taaagtaaca caagagttag gtcatgagga tcttatggct gcttatgtgg    6480

aaaacacaag cattaccatt aagaaaccta atgagctttc actagcctta ggtttaaaaa    6540

caattgccac tcatggtatt gctgcaatta atagtgttcc ttggagtaaa attttggctt    6600

atgtcaaacc attcttagga caagcagcaa ttacaacatc aaattgcgct aagagattag    6660

cacaacgtgt gtttaacaat tatatgcctt atgtgtttac attattgttc caattgtgta    6720

ctttttactaa aagtaccaat tctagaatta gagcttcact acctacaact attgctaaaa    6780

atagtgttaa gagtgttgct aaattatgtt tggatgccgg cattaattat gtgaagtcac    6840

ccaaatttc taaattgttc acaatcgcta tgtggctatt gttgttaagt atttgcttag    6900

gttctctaat ctgtgtaact gctgcttttg gtgtactctt atctaatttt ggtgctcctt    6960

cttattgtaa tggcgttaga gaattgtatc ttaattcgtc taacgttact actatggatt    7020

tctgtgaagg ttcttttcct tgcagcattt gtttaagtgg attagactcc cttgattctt    7080

atccagctct tgaaaccatt caggtgacga tttcatcgta caagctagac ttgacaattt    7140

taggtctggc cgctgagtgg gtttttggcat atatgttgtt cacaaaattc ttttatttat    7200

taggtctttc agctataatg caggtgttct ttggctattt tgctagtcat ttcatcagca    7260

attcttggct catgtggttt atcattagta ttgtacaaat ggcacccgtt tctgcaatgg    7320

ttaggatgta catcttcttt gcttctttct actacatatg gaagagctat gttcatatca    7380

tggatggttg cacctcttcg acttgcatga tgtgctataa gcgcaatcgt gccacacgcg    7440
```

```
ttgagtgtac aactattgtt aatggcatga agagatcttt ctatgtctat gcaaatggag    7500

gccgtggctt ctgcaagact cacaattgga attgtctcaa ttgtgacaca ttttgcactg    7560

gtagtacatt cattagtgat gaagttgctc gtgatttgtc actccagttt aaaagaccaa    7620

tcaaccctac tgaccagtca tcgtatattg ttgatagtgt tgctgtgaaa aatggcgcgc    7680

ttcacctcta ctttgacaag gctggtcaaa agacctatga gagacatccg ctctcccatt    7740

ttgtcaattt agacaatttg agagctaaca acactaaagg ttcactgcct attaatgtca    7800

tagtttttga tggcaagtcc aaatgcgacg agtctgcttc taagtctgct tctgtgtact    7860

acagtcagct gatgtgccaa cctattctgt tgcttgacca agctcttgta tcagacgttg    7920

gagatagtac tgaagtttcc gttaagatgt ttgatgctta tgtcgacacc ttttcagcaa    7980

cttttagtgt tcctatggaa aaacttaagg cacttgttgc tacagctcac agcgagttag    8040

caaagggtgt agctttagat ggtgtccttt ctacattcgt gtcagctgcc cgacaaggtg    8100

ttgttgatac cgatgttgac acaaaggatg ttattgaatg tctcaaactt tcacatcact    8160

ctgacttaga agtgacaggt gacagttgta acaatttcat gctcacctat aataaggttg    8220

aaaacatgac gcccagagat cttggcgcat gtattgactg taatgcaagg catatcaatg    8280

cccaagtagc aaaaagtcac aatgtttcac tcatctggaa tgtaaaagac tacatgtctt    8340

tatctgaaca gctgcgtaaa caaattcgta gtgctgccaa gaagaacaac atacctttta    8400

gactaacttg tgctacaact agacaggttg tcaatgtcat aactactaaa atctcactca    8460

agggtggtaa gattgttagt acttgtttta aacttatgct taaggccaca ttattgtgcg    8520

ttcttgctgc attggtttgt tatatcgtta tgccagtaca tacattgtca atccatgatg    8580

gttacacaaa tgaaatcatt ggttacaaag ccattcagga tggtgtcact cgtgacatca    8640

tttctactga tgattgtttt gcaaataaac atgctggttt tgacgcatgg tttagccagc    8700

gtggtggttc atacaaaaat gacaaaagct gccctgtagt agctgctatc attacaagag    8760

agattggttt catagtgcct ggcttaccgg tactgtgct gagagcaatc aatggtgact    8820

tcttgcattt tctacctcgt gtttttagtg ctgttggcaa catttgctac acaccttcca    8880

aactcattga gtatagtgat tttgctacct ctgcttgcgt tcttgctgct gagtgtacaa    8940

tttttaagga tgctatgggc aaacctgtgc catattgtta tgacactaat ttgctagagg    9000

gttctatttc ttatagtgag cttcgtccag acactcgtta tgtgcttatg gatggttcca    9060
```

```
tcatacagtt tcctaacact tacctggagg gttctgttag agtagtaaca acttttgatg    9120

ctgagtactg tagacatggt acatgcgaaa ggtcagaagt aggtatttgc ctatctacca    9180

gtggtagatg ggttcttaat aatgagcatt acagagctct atcaggagtt ttctgtggtg    9240

ttgatgcgat gaatctcata gctaacatct ttactcctct tgtgcaacct gtgggtgctt    9300

tagatgtgtc tgcttcagta gtggctggtg gtattattgc catattggtg acttgtgctg    9360

cctactactt tatgaaattc agacgtgttt ttggtgagta caaccatgtt gttgctgcta    9420

atgcactttt gttttttgatg tctttcacta tactctgtct ggtaccagct tacagctttc    9480

tgccgggagt ctactcagtc ttttacttgt acttgacatt ctatttcacc aatgatgttt    9540

cattcttggc tcaccttcaa tggtttgcca tgttttctcc tattgtgcct ttttggataa    9600

cagcaatcta tgtattctgt atttctctga agcactgcca ttggttcttt aacaactatc    9660

ttaggaaaag agtcatgttt aatggagtta catttagtac cttcgaggag gctgctttgt    9720

gtaccttttt gctcaacaag gaaatgtacc taaaattgcg tagcgagaca ctgttgccac    9780

ttacacagta taacaggtat cttgctctat ataacaagta caagtatttc agtggagcct    9840

tagatactac cagctatcgt gaagcagctt gctgccactt agcaaaggct ctaaatgact    9900

ttagcaactc aggtgctgat gttctctacc aaccaccaca gacatcaatc acttctgctg    9960

ttctgcagag tggttttagg aaaatggcat tcccgtcagg caaagttgaa gggtgcatgg   10020

tacaagtaac ctgtggaact acaactctta atggattgtg gttggatgac acagtatact   10080

gtccaagaca tgtcatttgc acagcagaag acatgcttaa tcctaactat gaagatctgc   10140

tcattcgcaa atccaaccat agctttcttg ttcaggctgg caatgttcaa cttcgtgtta   10200

ttggccattc tatgcaaaat tgtctgctta ggcttaaagt tgatacttct aaccctaaga   10260

cacccaagta taaatttgtc cgtatccaac ctggtcaaac attttcagtt ctagcatgct   10320

acaatggttc accatctggt gtttatcagt gtgccatgag acctaatcat accattaaag   10380

gttctttcct taatggatca tgtggtagtg ttggttttaa cattgattat gattgcgtgt   10440

ctttctgcta tatgcatcat atggagcttc caacaggagt acacgctggt actgacttag   10500

aaggtaaatt ctatggtcca tttgttgaca gacaaactgc acaggctgca ggtacagaca   10560

caaccataac attaaatgtt ttggcatggc tgtatgctgc tgttatcaat ggtgataggt   10620

ggtttcttaa tagattcacc actactttga atgactttaa ccttgtggca atgaagtaca   10680

actatgaacc tttgacacaa gatcatgttg acatattggg acctctttct gctcaaacag   10740
```

```
gaattgccgt cttagatatg tgtgctgctt tgaaagagct gctgcagaat ggtatgaatg   10800

gtcgtactat ccttggtagc actattttag aagatgagtt tacaccattt gatgttgtta   10860

gacaatgctc tggtgttacc ttccaaggta agttcaagaa aattgttaag ggcactcatc   10920

attggatgct tttaactttc ttgacatcac tattgattct tgttcaaagt acacagtggt   10980

cactgttttt ctttgtttac gagaatgctt tcttgccatt tactcttggt attatggcaa   11040

ttgctgcatg tgctatgctg cttgttaagc ataagcacgc attcttgtgc ttgtttctgt   11100

taccttctct tgcaacagtt gcttacttta atatggtcta catgcctgct agctgggtga   11160

tgcgtatcat gacatggctt gaattggctg acactagctt gtctggttat aggcttaagg   11220

attgtgttat gtatgcttca gctttagttt tgcttattct catgacagct cgcactgttt   11280

atgatgatgc tgctagacgt gtttggacac tgatgaatgt cattacactt gtttacaaag   11340

tctactatgg taatgcttta gatcaagcta tttccatgtg ggccttagtt atttctgtaa   11400

cctctaacta ttctggtgtc gttacgacta tcatgttttt agctagagct atagtgtttg   11460

tgtgtgttga gtattaccca ttgttattta ttactggcaa caccttacag tgtatcatgc   11520

ttgtttattg tttcttaggc tattgttgct gctgctactt tggccttttc tgtttactca   11580

accgttactt caggcttact cttggtgttt atgactactt ggtctctaca caagaattta   11640

ggtatatgaa ctcccagggg cttttgcctc ctaagagtag tattgatgct ttcaagctta   11700

acattaagtt gttgggtatt ggaggtaaac catgtatcaa ggttgctact gtacagtcta   11760

aaatgtctga cgtaaagtgc acatctgtgg tactgctctc ggttcttcaa caacttagag   11820

tagagtcatc ttctaaattg tgggcacaat gtgtacaact ccacaatgat attcttcttg   11880

caaaagacac aactgaagct ttcgagaaga tggtttctct tttgtctgtt ttgctatcca   11940

tgcagggtgc tgtagacatt aataggttgt gcgaggaaat gctcgataac cgtgctactc   12000

ttcaggctat tgcttcagaa tttagttctt taccatcata tgccgcttat gccactgccc   12060

aggaggccta tgagcaggct gtagctaatg gtgattctga agtcgttctc aaaaagttaa   12120

agaaatcttt gaatgtggct aaatctgagt ttgaccgtga tgctgccatg caacgcaagt   12180

tggaaaagat ggcagatcag gctatgaccc aaatgtacaa acaggcaaga tctgaggaca   12240

agaggcaaa agtaactagt gctatgcaaa caatgctctt cactatgctt aggaagcttg   12300

ataatgatgc acttaacaac attatcaaca atgcgcgtga tggttgtgtt ccactcaaca   12360
```

29

```
tcataccatt gactacagca gccaaactca tggttgttgt ccctgattat ggtacctaca   12420

agaacacttg tgatggtaac acctttacat atgcatctgc actctgggaa atccagcaag   12480

ttgttgatgc ggatagcaag attgttcaac ttagtgaaat taacatggac aattcaccaa   12540

atttggcttg gcctcttatt gttacagctc taagagccaa ctcagctgtt aaactacaga   12600

ataatgaact gagtccagta gcactacgac agatgtcctg tgcggctggt accacacaaa   12660

cagcttgtac tgatgacaat gcacttgcct actataacaa ttcgaaggga ggtaggtttg   12720

tgctggcatt actatcagac caccaagatc tcaaatgggc tagattccct aagagtgatg   12780

gtacaggtac aatttacaca gaactggaac caccttgtag gtttgttaca gacacaccaa   12840

aagggcctaa agtgaaatac ttgtacttca tcaaaggctt aaacaaccta aatagaggta   12900

tggtgctggg cagtttagct gctacagtac gtcttcaggc tggaaatgct acagaagtac   12960

ctgccaattc aactgtgctt tccttctgtg cttttgcagt agaccctgct aaagcatata   13020

aggattacct agcaagtgga ggacaaccaa tcaccaactg tgtgaagatg ttgtgtacac   13080

acactggtac aggacaggca attactgtaa caccagaagc taacatggac caagagtcct   13140

ttggtggtgc ttcatgttgt ctgtattgta gatgccacat tgaccatcca aatcctaaag   13200

gattctgtga cttgaaaggt aagtacgtcc aaatacctac cacttgtgct aatgacccag   13260

tgggttttac acttagaaac acagtctgta ccgtctgcgg aatgtggaaa ggttatggct   13320

gtagttgtga ccaactccgc gaacccttga tgcagtctgc ggatgcatca acgtttttaa   13380

acgggtttgc ggtgtaagtg cagcccgtct tacaccgtgc ggcacaggca ctagtactga   13440

tgtcgtctac agggcttttg atatttacaa cgaaaaagtt gctggttttg caaagttcct   13500

aaaaactaat tgctgtcgct tccaggagaa ggatgaggaa ggcaatttat tagactctta   13560

ctttgtagtt aagaggcata ctatgtctaa ctaccaacat gaagagacta tttataactt   13620

ggttaaagat tgtccagcgg ttgctgtcca tgactttttc aagtttagag tagatggtga   13680

catggtacca catatatcac gtcagcgtct aactaaatac acaatggctg atttagtcta   13740

tgctctacgt cattttgatg agggtaattg tgatacatta aaagaaatac tcgtcacata   13800

caattgctgt gatgatgatt atttcaataa gaaggattgg tatgacttcg tagagaatcc   13860

tgacatctta cgcgtatatg ctaacttagg tgagcgtgta cgccaatcat tattaaagac   13920

tgtacaattc tgcgatgcta tgcgtgatgc aggcattgta ggcgtactga cattagataa   13980

tcaggatctt aatgggaact ggtacgattt cggtgatttc gtacaagtag caccaggctg   14040
```

30

```
cggagttcct attgtggatt catattactc attgctgatg cccatcctca ctttgactag    14100

ggcattggct gctgagtccc atatggatgc tgatctcgca aaaccactta ttaagtggga    14160

tttgctgaaa tatgatttta cggaagagag actttgtctc ttcgaccgtt attttaaata    14220

ttgggaccag acataccatc ccaattgtat taactgtttg gatgataggt gtatccttca    14280

ttgtgcaaac tttaatgtgt tattttctac tgtgtttcca cctacaagtt ttggaccact    14340

agtaagaaaa atatttgtag atggtgttcc ttttgttgtt tcaactggat accattttcg    14400

tgagttagga gtcgtacata atcaggatgt aaacttacat agctcgcgtc tcagtttcaa    14460

ggaactttta gtgtatgctg ctgatccagc tatgcatgca gcttctggca atttattgct    14520

agataaacgc actacatgct tttcagtagc tgcactaaca aacaatgttg cttttcaaac    14580

tgtcaaaccc ggtaatttta ataaagactt ttatgacttt gctgtgtcta aaggtttctt    14640

taaggaagga agttctgttg aactaaaaca cttcttcttt gctcaggatg gcaacgctgc    14700

tatcagtgat tatgactatt atcgttataa tctgccaaca atgtgtgata tcagacaact    14760

cctattcgta gttgaagttg ttgataaata ctttgattgt tacgatggtg ctgtattaa    14820

tgccaaccaa gtaatcgtta acaatctgga taaatcagct ggtttcccat ttaataaatg    14880

gggtaaggct agactttatt atgactcaat gagttatgag gatcaagatg cacttttcgc    14940

gtatactaag cgtaatgtca tccctactat aactcaaatg aatcttaagt atgccattag    15000

tgcaaagaat agagctcgca ccgtagctgg tgtctctatc tgtagtacta tgacaaatag    15060

acagtttcat cagaaattat tgaagtcaat agccgccact agaggagcta ctgtggtaat    15120

tggaacaagc aagttttacg gtggctggca taatatgtta aaaactgttt acagtgatgt    15180

agaaactcca caccttatgg gttgggatta tccaaaatgt gacagagcca tgcctaacat    15240

gcttaggata atggcctctc ttgttcttgc tcgcaaacat aacacttgct gtaacttatc    15300

acaccgtttc tacaggttag ctaacgagtg tgcgcaagta ttaagtgaga tggtcatgtg    15360

tggcggctca ctatatgtta aaccaggtgg aacatcatcc ggtgatgcta caactgctta    15420

tgctaatagt gtctttaaca tttgtcaagc tgttacagcc aatgtaaatg cacttctttc    15480

aactgatggt aataagatag ctgacaagta tgtccgcaat ctacaacaca ggctctatga    15540

gtgtctctat agaaataggg atgttgatca tgaattcgtg gatgagtttt acgcttacct    15600

gcgtaaacat ttctccatga tgattctttc tgatgatgcc gttgtgtgct ataacagtaa    15660
```

31

```
ctatgcggct caaggtttag tagctagcat taagaacttt aaggcagttc tttattatca   15720

aaataatgtg ttcatgtctg aggcaaaatg ttggactgag actgacctta ctaaaggacc   15780

tcacgaattt tgctcacagc atacaatgct agttaaacaa ggagatgatt acgtgtacct   15840

gccttaccca gatccatcaa gaatattagg cgcaggctgt tttgtcgatg atattgtcaa   15900

aacagatggt acacttatga ttgaaaggtt cgtgtcactg gctattgatg cttacccact   15960

tacaaaacat cctaatcagg agtatgctga tgtctttcac ttgtatttac aatacattag   16020

aaagttacat gatgagctta ctggccacat gttggacatg tattccgtaa tgctaactaa   16080

tgataacacc tcacggtact gggaacctga gttttatgag ctatgtaca caccacatac   16140

agtcttgcag gctgtaggtg cttgtgtatt gtgcaattca cagacttcac ttcgttgcgg   16200

tgcctgtatt aggagaccat tcctatgttg caagtgctgc tatgaccatg tcatttcaac   16260

atcacacaaa ttagtgttgt ctgttaatcc ctatgtttgc aatgccccag gttgtgatgt   16320

cactgatgtg acacaactgt atctaggagg tatgagctat tattgcaagt cacataagcc   16380

tcccattagt tttccattat gtgctaatgg tcaggttttt ggtttataca aaaacacatg   16440

tgtaggcagt gacaatgtca ctgacttcaa tgcgatagca acatgtgatt ggactaatgc   16500

tggcgattac atacttgcca acacttgtac tgagagactc aagcttttcg cagcagaaac   16560

gctcaaagcc actgaggaaa catttaagct gtcatatggt attgccactg tacgcgaagt   16620

actctctgac agagaattgc atctttcatg ggaggttgga aaacctagac caccattgaa   16680

cagaaactat gtctttactg gttaccgtgt aactaaaaat agtaaagtac agattggaga   16740

gtacaccttt gaaaaaggtg actatggtga tgctgttgtg tacagaggta ctacgacata   16800

caagttgaat gttggtgatt actttgtgtt gacatctcac actgtaatgc cacttagtgc   16860

acctactcta gtgccacaag agcactatgt gagaattact ggcttgtacc caacactcaa   16920

catctcagat gagtttttcta gcaatgttgc aaattatcaa aaggtcggca tgcaaaagta   16980

ctctacactc caaggaccac ctggtactgg taagagtcat tttgccatcg acttgctct   17040

ctattaccca tctgctcgca tagtgtatac ggcatgctct catgcagctg ttgatgccct   17100

atgtgaaaag gcattaaaat atttgcccat agataaatgt agtagaatca tacctgcgcg   17160

tgcgcgcgta gagtgttttg ataaattcaa agtgaattca acactagaac agtatgtttt   17220

ctgcactgta aatgcattgc cagaaacaac tgctgacatt gtagtctttg atgaaatctc   17280

tatggctact aattatgact tgagtgttgt caatgctaga cttcgtgcaa aacactacgt   17340
```

32

```
ctatattggc gatcctgctc aattaccagc cccccgcaca ttgctgacta aaggcacact   17400

agaaccagaa tattttaatt cagtgtgcag acttatgaaa acaataggtc cagacatgtt   17460

ccttggaact tgtcgccgtt gtcctgctga aattgttgac actgtgagtg ctttagttta   17520

tgacaataag ctaaaagcac acaaggataa gtcagctcaa tgcttcaaaa tgttctacaa   17580

aggtgttatt acacatgatg tttcatctgc aatcaacaga cctcaaatag gcgttgtaag   17640

agaatttctt acacgcaatc ctgcttggag aaaagctgtt tttatctcac cttataattc   17700

acagaacgct gtagcttcaa aaatcttagg attgcctacg cagactgttg attcatcaca   17760

gggttctgaa tatgactatg tcatattcac acaaactact gaaacagcac actcttgtaa   17820

tgtcaaccgc ttcaatgtgg ctatcacaag ggcaaaaatt ggcattttgt gcataatgtc   17880

tgatagagat ctttatgaca aactgcaatt tacaagtcta gaaataccac gtcgcaatgt   17940

ggctacatta caagcagaaa atgtaactgg actttttaag gactgtagta agatcattac   18000

tggtcttcat cctacacagg cacctacaca cctcagcgtt gatataaagt tcaagactga   18060

aggattatgt gttgacatac caggcatacc aaaggacatg acctaccgta gactcatctc   18120

tatgatgggt ttcaaaatga attaccaagt caatggttac cctaatatgt ttatcacccg   18180

cgaagaagct attcgtcacg ttcgtgcgtg gattggcttt gatgtagagg ctgtcatgc   18240

aactagagat gctgtgggta ctaacctacc tctccagcta ggattttcta caggtgttaa   18300

cttagtagct gtaccgactg gttatgttga cactgaaaat aacacagaat tcaccagagt   18360

taatgcaaaa cctccaccag gtgaccagtt taaacatctt ataccactca tgtataaagg   18420

cttgccctgg aatgtagtgc gtattaagat agtacaaatg ctcagtgata cactgaaagg   18480

attgtcagac agagtcgtgt tcgtcctttg ggcgcatggc tttgagctta catcaatgaa   18540

gtactttgtc aagattggac ctgaaagaac gtgttgtctg tgtgacaaac gtgcaacttg   18600

cttttctact tcatcagata cttatgcctg ctggaatcat tctgtgggtt ttgactatgt   18660

ctataaccca tttatgattg atgttcagca gtggggcttt acgggtaacc ttcagagtaa   18720

ccatgaccaa cattgccagg tacatggaaa tgcacatgtg gctagttgtg atgctatcat   18780

gactagatgt ttagcagtcc atgagtgctt tgttaagcgc gttgattggt ctgttgaata   18840

ccctattata ggagatgaac tgagggttaa ttctgcttgc agaaaagtac aacacatggt   18900

tgtgaagtct gcattgcttg ctgataagtt tccagttctt catgacatag gaaatccaaa   18960
```

33

```
ggctatcaag tgtgtgcctc aggctgaagt agaatggaag ttctacgatg ctcagccatg   19020

tagtgacaaa gcttacaaaa tagaggaact cttctattct tatgctatac atcacgataa   19080

attcactgat ggtgtttgtt tgttttggaa ttgtaacgtt gatcgttacc cagccaatgc   19140

aattgtgtgt aggtttgaca caagagtctt gtcaaacttg aacttaccag gctgtgatgg   19200

tggtagtttg tatgtgaata agcatgcatt ccacactcca gctttcgata aaagtgcatt   19260

tactaattta aagcaattgc ctttctttta ctattctgat agtccttgtg agtctcatgg   19320

caaacaagta gtgtcggata ttgattatgt tccactcaaa tctgctacgt gtattacacg   19380

atgcaattta ggtggtgctg tttgcagaca ccatgcaaat gagtaccgac agtacttgga   19440

tgcatataat atgatgattt ctgctggatt tagcctatgg atttacaaac aatttgatac   19500

ttataacctg tggaatacat ttaccaggtt acagagttta gaaaatgtgg cttataatgt   19560

tgttaataaa ggacactttg atggacacgc cggcgaagca cctgtttcca tcattaataa   19620

tgctgtttac acaaaggtag atggtattga tgtggagatc tttgaaaata agacaacact   19680

tcctgttaat gttgcatttg agctttgggc taagcgtaac attaaaccag tgccagagat   19740

taagatactc aataatttgg gtgttgatat cgctgctaat actgtaatct gggactacaa   19800

aagagaagcc ccagcacatg tatctacaat aggtgtctgc acaatgactg acattgccaa   19860

gaaacctact gagagtgctt gttcttcact tactgtcttg tttgatggta gagtggaagg   19920

acaggtagac cttttttagaa acgcccgtaa tggtgtttta ataacagaag gttcagtcaa   19980

aggtctaaca ccttcaaagg gaccagcaca agctagcgtc aatggagtca cattaattgg   20040

agaatcagta aaaacacagt ttaactactt taagaaagta gacggcatta ttcaacagtt   20100

gcctgaaacc tactttactc agagcagaga cttagaggat tttaagccca gatcacaaat   20160

ggaaactgac tttctcgagc tcgctatgga tgaattcata cagcgatata agctcgaggg   20220

ctatgccttc gaacacatcg tttatggaga tttcagtcat ggacaacttg gcggtcttca   20280

tttaatgata ggcttagcca agcgctcaca agattcacca cttaaattag aggattttat   20340

ccctatggac agcacagtga aaaattactt cataacagat gcgcaaacag gttcatcaaa   20400

atgtgtgtgt tctgtgattg atcttttact tgatgacttt gtcgagataa taaagtcaca   20460

agatttgtca gtgatttcaa aagtggtcaa ggttacaatt gactatgctg aaatttcatt   20520

catgctttgg tgtaaggatg gacatgttga aaccttctac ccaaaactac aagcaagtca   20580

agcgtggcaa ccaggtgttg cgatgcctaa cttgtacaag atgcaaagaa tgcttcttga   20640
```

34

```
aaagtgtgac cttcagaatt atggtgaaaa tgctgttata ccaaaaggaa taatgatgaa    20700

tgtcgcaaag tatactcaac tgtgtcaata cttaaataca cttactttag ctgtaccsta    20760

caacatgaga gttattcact ttggtgctgg ctctgataaa ggagttgcac caggtacagc    20820

tgtgctcaga caatggttgc caactggcac actacttgtc gattcagatc ttaatgactt    20880

cgtctccgac gcagattcta ctttaattgg agactgtgca acagtacata cggctaataa    20940

atgggacctt attattagcg atatgtatga ccctaggacc aaacatgtga caaaagagaa    21000

tgactctaaa gaagggtttt tcacttatct gtgtggattt ataaagcaaa aactagccct    21060

gggtggttct atagctgtaa agataacaga gcattcttgg aatgctgacc tttacaagct    21120

tatgggccat ttctcatggt ggacagcttt tgttacaaat gtaaatgcat catcatcgga    21180

agcatttta attggggcta actatcttgg caagccgaag gaacaaattg atggctatac    21240

catgcatgct aactacattt tctggaggaa cacaaatcct atccagttgt cttcctattc    21300

actctttgac atgagcaaat ttcctcttaa attaagagga actgctgtaa tgtctcttaa    21360

ggagaatcaa atcaatgata tgatttattc tcttctggaa aaaggtaggc ttatcattag    21420

agaaaacaac agagttgtgg tttcaagtga tattcttgtt aacaactaaa cgaacatgtt    21480

tattttctta ttatttctta ctctcactag tggtagtgac cttgaccggt gcaccacttt    21540

tgatgatgtt caagctccta attacactca acatacttca tctatgaggg gggtttacta    21600

tcctgatgaa atttttagat cagacactct ttatttaact caggatttat ttcttccatt    21660

ttattctaat gttacagggt ttcatactat taatcatacg tttggcaacc ctgtcatacc    21720

ttttaaggat ggtatttatt ttgctgccac agagaaatca aatgttgtcc gtggttgggt    21780

ttttggttct accatgaaca acaagtcaca gtcggtgatt attattaaca attctactaa    21840

tgttgttata cgagcatgta ctttgaatt gtgtgacaac cctttctttg ctgtttctaa    21900

acccatgggt acacagacac atactatgat attcgataat gcatttaatt gcactttcga    21960

gtacatatct gatgcctttt cgcttgatgt ttcagaaaag tcaggtaatt ttaaacactt    22020

acgagagttt gtgtttaaaa ataaagatgg gtttctctat gtttataagg ctatcaacc    22080

tatagatgta gttcgtgatc taccttctgg ttttaacact ttgaaaccta tttttaagtt    22140

gcctcttggt attaacatta caaatttag agccattctt acagcctttt cacctgctca    22200

agacatttgg ggcacgtcag ctgcagccta tttttgttggc tatttaaagc caactacatt    22260
```

```
tatgctcaag tatgatgaaa atggtacaat cacagatgct gttgattgtt ctcaaaatcc   22320

acttgctgaa ctcaaatgct ctgttaagag ctttgagatt gacaaaggaa tttaccagac   22380

ctctaatttc agggttgttc cctcaggaga tgttgtgaga ttccctaata ttacaaactt   22440

gtgtcctttt ggagaggttt ttaatgctac taaattccct tctgtctatg catgggagag   22500

aaaaaaaatt tctaattgtg ttgctgatta ctctgtgctc tacaactcaa cattttttc    22560

aacctttaag tgctatggcg tttctgccac taagttgaat gatctttgct tctccaatgt   22620

ctatgcagat tcttttgtag tcaagggaga tgatgtaaga caaatagcgc caggacaaac   22680

tggtgttatt gctgattata attataaatt gccagatgat ttcatgggtt gtgtccttgc   22740

ttggaatact aggaacattg atgctacttc aactggtaat tataattata aatataggta   22800

tcttagacat ggcaagctta ggccctttga gagagacata tctaatgtgc ctttctcccc   22860

tgatggcaaa ccttgcaccc cacctgctct taattgttat tggccattaa atgattatgg   22920

tttttacacc actactggca ttggctacca accttacaga gttgtagtac tttcttttga   22980

acttttaaat gcaccggcca cggtttgtgg accaaaatta tccactgacc ttattaagaa   23040

ccagtgtgtc aattttaatt ttaatggact cactggtact ggtgtgttaa ctccttcttc   23100

aaagagattt caaccatttc aacaatttgg ccgtgatgtt tctgatttca ctgattccgt   23160

tcgagatcct aaaacatctg aaatattaga catttcacct tgctcttttg ggggtgtaag   23220

tgtaattaca cctggaacaa atgcttcatc tgaagttgct gttctatatc aagatgttaa   23280

ctgcactgat gtttctacag caattcatgc agatcaactc acaccagctt ggcgcatata   23340

ttctactgga aacaatgtat tccagactca agcaggctgt cttataggag ctgagcatgt   23400

cgacacttct tatgagtgcg acattcctat tggagctggc atttgtgcta gttaccatac   23460

agtttcttta ttacgtagta ctagccaaaa atctattgtg gcttatacta tgtctttagg   23520

tgctgatagt tcaattgctt actctaataa caccattgct atacctacta acttttcaat   23580

tagcattact acagaagtaa tgcctgtttc tatggctaaa acctccgtag attgtaatat   23640

gtacatctgc ggagattcta ctgaatgtgc taatttgctt ctccaatatg gtagcttttg   23700

cacacaacta aatcgtgcac tctcaggtat tgctgctgaa caggatcgca acacacgtga   23760

agtgttcgct caagttaaac aaatgtacaa aacccccaact ttgaaatatt ttggtggttt   23820

taattttttca caaatattac ctgaccctct aaagccaact aagaggtctt ttattgagga   23880

cttgctcttt aataaggtga cactcgctga tgctggcttc atgaagcaat atggcgaatg   23940
```

36

```
cctaggtgat attaatgcta gagatctcat ttgtgcgcag aagttcaatg gacttacagt   24000

gttgccacct ctgctcactg atgatatgat tgctgcctac actgctgctc tagttagtgg   24060

tactgccact gctggatgga catttggtgc tggcgctgct cttcaaatac cttttgctat   24120

gcaaatggca tataggttca atggcattgg agttacccaa aatgttctct atgagaacca   24180

aaaacaaatc gccaaccaat ttaacaaggc gattagtcaa attcaagaat cacttacaac   24240

aacatcaact gcattgggca agctgcaaga cgttgttaac cagaatgctc aagcattaaa   24300

cacacttgtt aaacaactta gctctaattt tggtgcaatt tcaagtgtgc taaatgatat   24360

cctttcgcga cttgataaag tcgaggcgga ggtacaaatt gacaggttaa ttacaggcag   24420

acttcaaagc cttcaaacct atgtaacaca acaactaatc agggctgctg aaatcagggc   24480

ttctgctaat cttgctgcta ctaaaatgtc tgagtgtgtt cttggacaat caaaaagagt   24540

tgactttgt ggaaagggct accaccttat gtccttccca caagcagccc cgcatggtgt   24600

tgtcttccta catgtcacgt atgtgccatc ccaggagagg aacttcacca cagcgccagc   24660

aatttgtcat gaaggcaaag catacttccc tcgtgaaggt gtttttgtgt ttaatggcac   24720

ttcttggttt attacacaga ggaacttctt ttctccacaa ataattacta cagacaatac   24780

atttgtctca ggaaattgtg atgtcgttat tggcatcatt aacaacacag tttatgatcc   24840

tctgcaacct gagcttgact cattcaaaga agagctggac aagtacttca aaaatcatac   24900

atcaccagat gttgatcttg gcgacatttc aggcattaac gcttctgtcg tcaacattca   24960

aaaagaaatt gaccgcctca atgaggtcgc taaaaattta aatgaatcac tcattgacct   25020

tcaagaattg ggaaaatatg agcaatatat taaatggcct tggtatgttt ggctcggctt   25080

cattgctgga ctaattgcca tcgtcatggt tacaatcttg ctttgttgca tgactagttg   25140

ttgcagttgc ctcaagggtg catgctcttg tggttcttgc tgcaagtttg atgaggatga   25200

ctctgagcca gttctcaagg gtgtcaaatt acattacaca taaacgaact tatggatttg   25260

tttatgagat tttttactct tggatcaatt actgcacagc cagtaaaaat tgacaatgct   25320

tctcctgcaa gtactgttca tgctacagca acgataccgc tacaagcctc actccctttc   25380

ggatggcttg ttattggcgt tgcatttctt gctgttttтc agagcgctac caaaataatt   25440

gcgctcaata aaagatggca gctagccctt tataagggct tccagttcat ttgcaatttа   25500

ctgctgctat ttgttaccat ctattcacat cttttgcttg tcgctgcagg tatggaggcg   25560
```

37

```
caattttttgt  acctctatgc  cttgatatat  tttctacaat  gcatcaacgc  atgtagaatt  25620

attatgagat  gttggcttg   ttggaagtgc  aaatccaaga  acccattact  ttatgatgcc  25680

aactactttg  tttgctggca  cacacataac  tatgactact  gtataccata  taacagtgtc  25740

acagatacaa  ttgtcgttac  tgaaggtgac  ggcatttcaa  caccaaaact  caaagaagac  25800

taccaaattg  gtggttattc  tgaggatagg  cactcaggtg  ttaaagacta  tgtcgttgta  25860

catggctatt  tcaccgaagt  ttactaccag  cttgagtcta  cacaaattac  tacagacact  25920

ggtattgaaa  atgctacatt  cttcatcttt  aacaagcttg  ttaaagaccc  accgaatgtg  25980

caaatacaca  caatcgacgg  ctcttcagga  gttgctaatc  cagcaatgga  tccaatttat  26040

gatgagccga  cgacgactac  tagcgtgcct  ttgtaagcac  aagaaagtga  gtacgaactt  26100

atgtactcat  tcgtttcgga  agaaacaggt  acgttaatag  ttaatagcgt  acttcttttt  26160

cttgctttcg  tggtattctt  gctagtcaca  ctagccatcc  ttactgcgct  tcgattgtgt  26220

gcgtactgct  gcaatattgt  taacgtgagt  ttagtaaaac  caacggttta  cgtctactcg  26280

cgtgttaaaa  atctgaactc  ttctgaagga  gttcctgatc  ttctggtcta  aacgaactaa  26340

ctattattat  tattctgttt  ggaactttaa  cattgcttat  catggcagac  aacggtacta  26400

ttaccgttga  ggagcttaaa  caactcctgg  aacaatggaa  cctagtaata  ggtttcctat  26460

tcctagcctg  gattatgtta  ctacaatttg  cctattctaa  tcggaacagg  tttttgtaca  26520

taataaagct  tgttttcctc  tggctcttgt  ggccagtaac  acttgcttgt  tttgtgcttg  26580

ctgctgtcta  cagaattaat  tgggtgactg  gcgggattgc  gattgcaatg  gcttgtattg  26640

taggcttgat  gtggcttagc  tacttcgttg  cttccttcag  gctgtttgct  cgtacccgct  26700

caatgtggtc  attcaaccca  gaaacaaaca  ttcttctcaa  tgtgcctctc  cggggggacaa  26760

ttgtgaccag  accgctcatg  gaaagtgaac  ttgtcattgg  tgctgtgatc  attcgtggtc  26820

acttgcgaat  ggccggacac  tccctagggc  gctgtgacat  taaggacctg  ccaaaagaga  26880

tcactgtggc  tacatcacga  acgctttctt  attacaaatt  aggagcgtcg  cagcgtgtag  26940

gcactgattc  aggttttgct  gcatacaacc  gctaccgtat  tggaaactat  aaattaaata  27000

cagaccacgc  cggtagcaac  gacaatattg  ctttgctagt  acagtaagtg  acaacagatg  27060

tttcatcttg  ttgacttcca  ggttacaata  gcagagatat  tgattatcat  tatgaggact  27120

ttcaggattg  ctatttggaa  tcttgacgtt  ataataagtt  caatagtgag  acaattattt  27180

aagcctctaa  ctaagaagaa  ttattcggag  ttagatgatg  aagaacctat  ggagttagat  27240
```

```
tatccataaa  acgaacatga  aaattattct  cttcctgaca  ttgattgtat  ttacatcttg  27300

cgagctatat  cactatcagg  agtgtgttag  aggtacgact  gtactactaa  aagaaccttg  27360

cccatcagga  acatacgagg  gcaattcacc  atttcaccct  cttgctgaca  ataaatttgc  27420

actaacttgc  actagcacac  actttgcttt  tgcttgtgct  gacggtactc  gacataccta  27480

tcagctgcgt  gcaagatcag  tttcaccaaa  acttttcatc  agacaagagg  aggttcaaca  27540

agagctctac  tcgccacttt  ttctcattgt  tgctgctcta  gtatttttaa  tactttgctt  27600

caccattaag  agaaagacag  aatgaatgag  ctcactttaa  ttgacttcta  tttgtgcttt  27660

ttagcctttc  tgctattcct  tgttttaata  atgcttatta  tattttggtt  ttcactcgaa  27720

atccaggatc  tagaagaacc  ttgtaccaaa  gtctaaacga  acatgaaact  tctcattgtt  27780

ttgacttgta  tttctctatg  cagttgcata  tgcactgtag  tacagcgctg  tgcatctaat  27840

aaacctcatg  tgcttgaaga  tccttgtaag  gtacaacact  aggggtaata  cttatagcac  27900

tgcttggctt  tgtgctctag  gaaaggtttt  acctttttcat  agatggcaca  ctatggttca  27960

aacatgcaca  cctaatgtta  ctatcaactg  tcaagatcca  gctggtggtg  cgcttatagc  28020

taggtgttgg  taccttcatg  aaggtcacca  aactgctgca  tttagagacg  tacttgttgt  28080

tttaaataaa  cgaacaaatt  aaaatgtctg  ataatggacc  ccaatcaaac  caacgtagtg  28140

cccccgcat   tacatttggt  ggacccacag  attcaactga  caataaccag  aatggaggac  28200

gcaatggggc  aaggccaaaa  cagcgccgac  cccaaggttt  acccaataat  actgcgtctt  28260

ggttcacagc  tctcactcag  catggcaagg  aggaacttag  attccctcga  ggccagggcg  28320

ttccaatcaa  caccaatagt  ggtccagatg  accaaattgg  ctactaccga  agagctaccc  28380

gacgagttcg  tggtggtgac  ggcaaaatga  aagagctcag  ccccagatgg  tacttctatt  28440

acctaggaac  tggcccagaa  gcttcacttc  cctacggcgc  taacaaagaa  ggcatcgtat  28500

gggttgcaac  tgagggagcc  ttgaatacac  ccaaagacca  cattggcacc  cgcaatccta  28560

ataacaatgc  tgccaccgtg  ctacaacttc  ctcaaggaac  aacattgcca  aaaggcttct  28620

acgcagaggg  aagcagaggc  ggcagtcaag  cctcttctcg  ctcctcatca  cgtagtcgcg  28680

gtaattcaag  aaattcaact  cctggcagca  gtaggggaaa  ttctcctgct  cgaatggcta  28740

gcggaggtgg  tgaaactgcc  ctcgcgctat  tgctgctaga  cagattgaac  cagcttgaga  28800

gcaaagtttc  tggtaaaggc  caacaacaac  aaggccaaac  tgtcactaag  aaatctgctg  28860
```

39

```
ctgaggcatc taaaaagcct cgccaaaaac gtactgccac aaaacagtac aacgtcactc    28920

aagcatttgg gagacgtggt ccagaacaaa cccaaggaaa tttcggggac caagacctaa    28980

tcagacaagg aactgattac aaacattggc cgcaaattgc acaatttgct ccaagtgcct    29040

ctgcattctt tggaatgtca cgcattggca tggaagtcac accttcggga acatggctga    29100

cttatcatgg agccattaaa ttggatgaca aagatccaca attcaaagac aacgtcatac    29160

tgctgaacaa gcacattgac gcatacaaaa cattcccacc aacagagcct aaaaaggaca    29220

aaaagaaaaa gactgatgaa gctcagcctt gccgcagag acaaaagaag cagcccactg    29280

tgactcttct tcctgcggct gacatggatg atttctccag acaacttcaa aattccatga    29340

gtggagcttc tgctgattca actcaggcat aaacactcat gatgaccaca caaggcagat    29400

gggctatgta aacgttttcg caattccgtt tacgatacat agtctactct tgtgcagaat    29460

gaattctcgt aactaaacag cacaagtagg tttagttaac tttaatctca catagcaatc    29520

tttaatcaat gtgtaacatt agggaggact tgaaagagcc accacatttt catcgaggcc    29580

acgcggagta cgatcgaggg tacagtgaat aatgctaggg agagctgcct atatggaaga    29640

gccctaatgt gtaaaattaa ttttagtagt gctatcccca tgtgattta atagcttctt    29700

aggagaatga c                                                         29711


<210>   10
<211>   31
<212>   DNA
<213>   SARS coronavirus

<400>   10
cgggatccat gtctgataat ggacccccaat c                                          31


<210>   11
<211>   31
<212>   DNA
<213>   SARS coronavirus

<400>   11
acgcgtcgac ttatgcctga gttgaatcag c                                           31


<210>   12
<211>   31
<212>   DNA
<213>   SARS coronavirus

<400>   12
```

cgggatccat gtctgataat ggaccccaat c   31

<210>   13
<211>   30
<212>   DNA
<213>   SARS coronavirus

<400>   13
acgcgtcgac tcgagcagga gaatttcccc   30

<210>   14
<211>   31
<212>   DNA
<213>   SARS coronavirus

<400>   14
cgggatccaa ccagcttgag agcaaagttt c   31

<210>   15
<211>   31
<212>   DNA
<213>   SARS coronavirus

<400>   15
acgcgtcgac ttatgcctga gttgaatcag c   31

<210>   16
<211>   29
<212>   DNA
<213>   SARS coronavirus

<400>   16
cgggatccgc cttgaataca cccaaagac   29

<210>   17
<211>   30
<212>   DNA
<213>   SARS coronavirus

<400>   17
acgcgtcgac aaattgtgca atttgcggcc   30

<210>   18
<211>   29
<212>   DNA
<213>   SARS coronavirus

<400>   18
cgggatccgc cttgaataca cccaaagac   29

41

```
<210>    19
<211>    28
<212>    DNA
<213>    SARS coronavirus

<400>    19
acgcgtcgac agcaggagaa tttcccct                                    28


<210>    20
<211>    29
<212>    DNA
<213>    SARS coronavirus

<400>    20
cgggatcctt gaaccagctt gagagcaaa                                   29


<210>    21
<211>    30
<212>    DNA
<213>    SARS coronavirus

<400>    21
acgcgtcgac aaattgtgca atttgcggcc                                  30


<210>    22
<211>    29
<212>    DNA
<213>    SARS coronavirus

<400>    22
cgggatccga tccacaattc aaagacaac                                   29


<210>    23
<211>    31
<212>    DNA
<213>    SARS coronavirus

<400>    23
acgcgtcgac ttatgcctga gttgaatcag c                                31


<210>    24
<211>    32
<212>    DNA
<213>    SARS coronavirus


<220>
<221>    misc_feature
```

<222> (3)..(8)

<400> 24
cgggatccaa cgtcatactg ctgaacaagc ac   32

<210> 25
<211> 31
<212> DNA
<213> SARS coronavirus

<220>
<221> misc_feature
<222> (5)..(10)

<400> 25
acgcgtcgac ttatgcctga gttgaatcag c   31

43